**SUPPLEMENTAL MATERIAL**

**Predicting chemically-induced skin reactions. Part II: QSAR Models of Skin Permeability and the Relationship between Skin Permeability and Skin Sensitization**

Vinicius M. Alves,a,b Eugene Muratov,b,c Denis Fourches,b Judy Strickland,d Nicole Kleinstreuer,d Carolina H. Andrade,a and Alexander Tropshab,\*.

a Laboratory of Molecular Modeling and Design, Faculty of Pharmacy, Federal University of Goiás, Goiânia, GO, 74605-220, Brazil.

b Laboratory for Molecular Modeling, Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, 27599, USA.

c Laboratory of Theoretical Chemistry, A.V. Bogatsky Physical-Chemical Institute NAS of Ukraine, Odessa, 65080, Ukraine.

d ILS, Inc./Contractor supporting the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), P.O. Box 13501, Research Triangle Park, NC, 27709, USA.

**Corresponding Author**

\* Address for correspondence: 100K Beard Hall, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, 27599, USA; Telephone: (919) 966-2955; FAX: (919) 966-0204; E-mail: [alex\_tropsha@unc.edu](mailto:alex_tropsha@unc.edu)

**TABLE OF CONTENTS**

[**Table S1.** Previous QSAR studies of skin permeability. 3](#_Toc378679961)

[**Table S2.** Pairs of the structural duplicates found in the dataset used by Chauhan and Shakya23. 5](#_Toc378679962)

[**Table S3.** Examples of "suspicious" compounds in human permeability dataset D identified by cluster analysis. 6](#_Toc378679963)

[**Table S4.** United skin sensitization and human permeability dataset with imputed data. 8](#_Toc378679964)

[**Table S5.** Compounds within Cluster 1 of united skin sensitization and permeability dataset. 16](#_Toc378679965)

[**Figure S1.** Skin permeability *vs.* sensitization imputed data for united dataset, containing 174 sensitizers and 109 non-sensitizers. 20](#_Toc378679966)

[**Figure S2.** Skin permeability *vs.* sensitization imputed data for the Cluster 1 of united dataset, containing 4 sensitizers and 7 non-sensitizers. 21](#_Toc378679967)

[**REFERENCES** 22](#_Toc378679968)

# STATISTICAL METRICS

Cross-validation correlation coefficient ():

 Equation (1)

Root Mean Squared Error (RMSE):

 Equation (2)

Mean Absolute Error (MAE):

 Equation (3)

Cross-validation correlation coefficient 3 ()1,2:

 Equation (4)

metrics3–6:

 Equation (5)

 Equation (6)

 Equation (7)

where *n* is the total number of compounds, ***Y****i,exp* is the experimental activity of i*th* molecule, ***Y****i,pred* is the predicted activity of i*th*molecule, <***Y****>exp* is the averaged experimental activity over all *n* compounds, <***Y****>tr*is the averaged experimental data values in the training set, *n*ext the number of compounds in the external set (folds), and *ntr* the number of compounds in the training set, *r2* and are respectively the determination coefficients of the regression function calculated using the experimental and predicted data of the external set.

# Table S1. Previous QSAR studies of skin permeability.

| **Type of Assay/Endpoint** | **Number of compounds** | **Dataset source** | **Descriptors** | **Statistical Methods** | **Validation** | **Models** | **AD** | **Y-rand** | **Ref.** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Human skin  LogKp (cm/h) | 93 | 7 | MV, MW, Kow | MRA | Internal | r2=0.670 | No | No | 8 |
| Human skin  LogKp (cm/h) | 31 | 9–11 | LogP, MW, Hb | SR | External | *q*2=0.960 | No | No | 12 |
| Human skin  LogKp (cm/h) | 60 | 7 | MV, logP, mpt | MRA  PCA | Internal | r2=0.885 | No | No | 13 |
| Human skin  LogKp (cm/s) | 37 | 7 | MV, logP, logD0/δ, Hd, Ha, β | MRA | No | r2=0.940 | No | No | 14 |
| Human skin  LogKp (cm/s) | 28 | 7 | logP, MR, Hb, Ha, Vehicle, Occlusion | MRA | No | r2=0.893 | No | No | 15 |
| Human skin  LogKp (cm/h) | 53 | 7,16 | Abraham descriptors | MRA | No | r2=0.957 | No | No | 17 |
| Human skin  LogKp (cm/h) | 107 | 18 | Physicochemicals, topologicals, eletronics | SRA/LSRA | No | r=0.927 | No | No | 19 |
| Human skin  LogKp (cm/s) | 116 | 7,16,18,20 | LogP, MW | Regression Analysis | LOO | r2=0.810 | No | No | 21 |
| Human skin  LogKp (cm/h) | 143 | 7,22 | log*Kow*, VAMP, TSAR, QSARis, AMSOL, Dragon, Linearity index, log*K*dmpc-w | SRA | No | r2=0.900 | No | No | 23 |
| Human skin  LogJmax (mol/cm2/h) | 178 | Misc. | MW | Stepwise multivariate regression analysis | External | r2int=0.847  r2ext=0.784  r2ext=0.537  r2ext=0.282 r2ext=0.537 | No | No | 24 |
| Human skin  LogKp (cm/h) | 182 | 25,26 | log*Kow*, MW, water solubility | MRA | External | Residual variance = 0.47 | No | No | 27 |
| Human skin  LogKp (cm/h) | 215 | 7,23,28 | ABSOLV | ANN | External | *q*2ext=0.698 | No | No | 29 |
| MLR | *q*2ext=0.792 |
| Human skin  LogKp (cm/h) | >200 | Misc. | Brick and Mortar | | External | *q*2=0.74 | No | No | 30 |
| Human skin LogKp (cm/h) | 208 | 29–31 | e-Dragon, Pharma Algorithms ADME Tox web | MLR | External | *q*²= 0.932 | Yes | No | 32 |
| PCR | *q*²= 0.886 |
| PLSR | *q*2= 0.936 |
| Human Skin logKp (cm/h) | 140 | Misc. | logP, MW, Hd, Ha, solubility | GPR | External | MSE=1.13 | No | No | 33 |
| Rodent skin logKp (cm/h) | 103 | Misc. | MSE=0.53 |
| Pig skin logKp (cm/h) | 15 | Misc. | MSE=0.83 |
| Synthetic membranes  logKp (cm/h) | 19 | Misc. | MSE=3.57 |
| Pig skin logKp (cm/h) | 36 | In house | Several physical chemical parameters | MLR, SRA | External | 0.41-0.81 | Yes | No | 34 |
| Human Skin logKp (cm/h) | 283 | Misc. | Dragon, WHIM, GETAWAY | MPSO-MLR, ANFIS | External | q2=0.874  q2=0.890 | Yes | Yes | 35 |

Notes: AD: Applicability domain; ADME: absorption, distribution, metabolism, excretion; ANFIS: Adaptive neuro-fuzzy inference system; ANN:D: diffusivity through the membrane; GPR: Gaussian process regression; Ha: hydrogen bond acceptor; Hb: hydrogen bond-forming ability; Hd: hydrogen bond donor; Kow or P: Octanol/water partition; Kp: permeability coefficient; LSRA: Least-square regression analysis; log*K*dmpc-w: Logarithm of the L-α-DMPC-water partition coefficient; LOO: leave-one-out; Misc: Miscellaneous; MLR: Multiple Linear Regression; MPSO: modified particle swarm optimization; Mpt: Melting point; MRA: Multivariate regression analysis; MSE: Mean square error; MV: Molecular volume; MW: Molecular weight; PCA: Principal component analysis; PCR: Principal component regression; PLSR: Partial least-squares regression; SRA: Stepwise regression analysis; SVM: Support vector machine; Y-rand: Y-randomization; β: diffusivity constant; δ: Diffusion path length.

­

# Table S2. Statistical characteristics of QSAR models for skin permeability assessed by 5-fold external cross-validation.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Models** |  | **RMSE** | **MAE** |  |  |  |  | **Coverage** |
| H  U  M  A  N | Model 1 - SiRMS | 0.69 | 0.52 | 0.40 | 0.75 | 0.61 | 0.46 | 0.30 | 0.72 |
|  | Model 2 - SiRMS no AD\* | 0.50 | 0.73 | 0.53 | 0.50 | 0.48 | 0.32 | 0.34 | 1.00 |
|  | Model 3 - Dragon | 0.73 | 0.46 | 0.35 | 0.80 | 0.64 | 0.50 | 0.27 | 0.68 |
|  | Model 4 - Dragon no AD\* | 0.55 | 0.70 | 0.49 | 0.55 | 0.51 | 0.34 | 0.35 | 1.00 |
|  | Model 5 - Consensus | 0.72 | 0.49 | 0.38 | 0.77 | 0.62 | 0.48 | 0.28 | 0.77 |
|  | Model 6 - Consensus no AD\* | 0.55 | 0.69 | 0.50 | 0.55 | 0.51 | 0.33 | 0.36 | 1.00 |
|  | Model 7 - Consensus Rigor | 0.73 | 0.45 | 0.34 | 0.81 | 0.62 | 0.48 | 0.29 | 0.63 |
| R  O  D  E  N  T | Model 8 - SiRMS | 0.57 | 0.67 | 0.50 | 0.57 | 0.54 | 0.37 | 0.34 | 0.71 |
|  | Model 9 - SiRMS no AD\* | 0.35 | 0.82 | 0.61 | 0.35 | 0.34 | 0.18 | 0.31 | 1.00 |
|  | Model 10 - Dragon | 0.33 | 0.75 | 0.56 | 0.45 | 0.31 | 0.16 | 0.30 | 0.69 |
|  | Model 11 - Dragon no AD\* | 0.32 | 0.84 | 0.63 | 0.32 | 0.31 | 0.11 | 0.40 | 1.00 |
|  | Model 12 - Consensus | 0.41 | 0.77 | 0.56 | 0.43 | 0.41 | 0.36 | 0.36 | 0.86 |
|  | Model 13 - Consensus no AD\* | 0.38 | 0.80 | 0.61 | 0.38 | 0.35 | 0.15 | 0.41 | 1.00 |
|  | Model 14 - Consensus Rigor | 0.61 | 0.58 | 0.43 | 0.67 | 0.56 | 0.39 | 0.34 | 0.53 |

Notes: Models 1 to 7: Human-based skin permeability models. Models 8 to 13: Rodent-based skin permeability models. RSME: root mean square error; MAE: mean absolute error \*Applicability Domain was not considered in these models.

# Table S3. Pairs of the structural duplicates found in the dataset used by Chauhan and Shakya32.

|  |  |  |  |
| --- | --- | --- | --- |
| **Compound Name** | **log Kp** | **Duplicates** | **log Kp** |
| p-Phenylenediamine | -3.62 | 4-Phenylenediamine | -3.62 |
| Cortexone | -3.35 | Deoxycorticosterone | -3.34 |
| 2-Phenylenediamine | -3.35 | o-Phenylenediamine | -3.35 |
| Propanol | -2.85 | n-Propanol | -2.85 |
| Butanoic acid | -2.58 | Butyric acid | -3.00 |
| Lidocaine | -2.40 | Lignocaine | -2.40 |
| Butanone | -2.35 | 2-Butanone | -2.35 |
| p-Nitrophenol | -2.25 | 4-Nitrophenol | -2.25 |
| 3-Nitrophenol | -2.25 | m-Nitrophenol | -2.25 |
| 2-Methylphenol | -1.81 | o-Cresol | -1.81 |
| 3-Methylphenol | -1.81 | 3-Cresol; m-Cresol | -1.82 |
| Diethylether | -1.79 | Ethyl ether | -1.80 |
| 4-Methylphenol | -1.75 | 4-Cresol; p-Cresol | -1.76 |
| 2-Naphthol | -1.55 | β-Naphthol | -1.56 |
| 2-Chlorophenol | -1.48 | o-Chlorophenol | -1.48 |
| 4-ethylphenol | -1.46 | p-ethylphenol | -1.45 |
| 3,4-Xylenol | -1.44 | 3,4-Dimethylphenol | -1.44 |
| Thymol | -1.26 | 2-Isopropyl-5-methylphenol | -1.27 |
| 4-Chloro-3,5-dimethylphenol | -1.23 | 4-Chloro-3,5-xylenol | -1.28 |

Notes: Kp: permeability coefficient

# Table S4. Examples of "suspicious" compounds in human permeability dataset D identified by cluster analysis.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Compounds/Toxicity profiles** | | | |
| **Hydrocortisone derivatives** |  | |  | |
| Hydrocortisone hexanoate  logKp = -1.74 | | Hydrocortisone octanoate  logKp = -1.21 | |
|  | |  | |
| Hydrocortisone methylpimelate  logKp = -2.27 | | Hydrocortisone hemipimelate  logKp = -2.74 | |
| **Carboxylic acids** |  |  | |  |
| Acetic acid  logKp = -3.21 | Propionic acid  logKp = -2.74 | | Butanoic acid  logKp = -2.90 |
|  | |  | |
| Pentanoic acid  logKp = -2.70 | | Hexanoic acid  logKp = -1.85 | |
|  |  | |  | |
|  | Heptanoic acid  logKp = -1.70 | | Octanoic acid  logKp = -1.60 | |

Notes: Kp: permeability coefficient

# Table S5. United skin sensitization and human permeability dataset with imputed data.

|  | **Compound Name** | **logKp** | **Predicted Sensitization** | **Predicted logKp** | **LLNA Result** |
| --- | --- | --- | --- | --- | --- |
| 1 | 1-Dodecyl glycidyl ether | -5.48 | Sensitizer |  |  |
| 2 | Atropine | -5.07 | Non-sensitizer |  |  |
| 3 | Cortisone | -5.00 | Non-sensitizer |  |  |
| 4 | Mannitol | -4.96 | Non-sensitizer |  |  |
| 5 | 5-Fluorouracil | -4.78 | Non-sensitizer |  |  |
| 6 | Hydrocortisone succinamate | -4.59 | Non-sensitizer |  |  |
| 7 | 2,3-Butanediol | -4.40 | Non-sensitizer |  |  |
| 8 | Scopolamine | -4.30 | Non-sensitizer |  |  |
| 9 | Dexamethasone | -4.19 | Non-sensitizer |  |  |
| 10 | Cortexolone | -4.13 | Non-sensitizer |  |  |
| 11 | Ethylene glycol | -4.07 | Non-sensitizer |  |  |
| 12 | Ethanol amine | -4.02 | Non-sensitizer |  |  |
| 13 | Corticosterone | -4.00 | Non-sensitizer |  |  |
| 14 | Chlorpyrifos | -3.96 | Non-sensitizer |  |  |
| 15 | Barbital | -3.95 | Non-sensitizer |  |  |
| 16 | Hydrocortisone | -3.93 | Non-sensitizer |  |  |
| 17 | 1,6-Bis(2,3-epoxypropoxy)-hexane | -3.87 |  |  | Sensitizer |
| 18 | Urea | -3.83 | Non-sensitizer |  |  |
| 19 | n-Propoxyethanol | -3.77 | Sensitizer |  |  |
| 20 | Butobarbital | -3.71 | Non-sensitizer |  |  |
| 21 | Hydrocortisone methylsuccinate | -3.68 | Non-sensitizer |  |  |
| 22 | 4-Phenylene-diamine | -3.62 |  |  | Sensitizer |
| 23 | Resorcinol | -3.62 |  |  | Sensitizer |
| 24 | 2-Ethoxyethanol | -3.60 | Non-sensitizer |  |  |
| 25 | Nicorandil | -3.58 | Non-sensitizer |  |  |
| 26 | 2-Isopropoxyethanol | -3.57 | Non-sensitizer |  |  |
| 27 | Estradiol | -3.52 | Sensitizer |  |  |
| 28 | 2-Phenylenediamine | -3.35 | Sensitizer |  |  |
| 29 | Phenobarbital | -3.34 | Non-sensitizer |  |  |
| 30 | Methanol | -3.30 | Non-sensitizer |  |  |
| 31 | 2-Nitro-p-phenylenediamine | -3.30 |  |  | Sensitizer |
| 32 | Acetone | -3.29 | Non-sensitizer |  |  |
| 33 | Acetonitrile | -3.21 | Non-sensitizer |  |  |
| 34 | Acetaldehyde | -3.15 | Non-sensitizer |  |  |
| 35 | Ethanol | -3.10 | Non-sensitizer |  |  |
| 36 | Ethylamine | -3.09 | Non-sensitizer |  |  |
| 37 | Propylene oxide | -3.05 | Non-sensitizer |  |  |
| 38 | Acrylic acid | -3.05 | Sensitizer |  |  |
| 39 | Isopropanol | -3.05 |  |  | Non-sensitizer |
| 40 | 1-Methoxypropan-2-ol | -2.90 | Non-sensitizer |  |  |
| 41 | Butanoic acid | -2.90 | Non-sensitizer |  |  |
| 42 | Isopropylamine | -2.90 | Non-sensitizer |  |  |
| 43 | Griseofulvin | -2.89 | Non-sensitizer |  |  |
| 44 | 2-Phenoxyethanol | -2.88 | Sensitizer |  |  |
| 45 | Propanol | -2.85 | Non-sensitizer |  |  |
| 46 | 2-Butoxyethanol | -2.85 | Sensitizer |  |  |
| 47 | Fluocinonide | -2.77 | Non-sensitizer |  |  |
| 48 | Catechol | -2.77 | Sensitizer |  |  |
| 49 | Diethylamine | -2.75 | Non-sensitizer |  |  |
| 50 | Hydrocortisone hemipimelate | -2.74 | Non-sensitizer |  |  |
| 51 | Pyridine | -2.74 |  |  | Sensitizer |
| 52 | Pentanoic acid | -2.70 | Non-sensitizer |  |  |
| 53 | Clotrimazole | -2.70 |  |  | Sensitizer |
| 54 | Methyl acrylate | -2.68 |  |  | Sensitizer |
| 55 | Formaldehyde | -2.65 | Non-sensitizer |  |  |
| 56 | Isobutyl alcohol | -2.65 | Non-sensitizer |  |  |
| 57 | Aniline | -2.65 |  |  | Sensitizer |
| 58 | Amylobarbital | -2.64 | Non-sensitizer |  |  |
| 59 | Butanol | -2.60 |  |  | Non-sensitizer |
| 60 | 2-Methyl propenoic acid | -2.58 |  |  | Sensitizer |
| 61 | 4-Amino-2-nitrophenol | -2.55 | Sensitizer |  |  |
| 62 | Ethyl acrylate | -2.39 |  |  | Sensitizer |
| 63 | 1-1-1-Trichloroethane | -2.35 | Non-sensitizer |  |  |
| 64 | 2-Hexanone | -2.35 | Sensitizer |  |  |
| 65 | Butanone | -2.34 | Non-sensitizer |  |  |
| 66 | Hydrocortisone methylpimelate | -2.27 | Non-sensitizer |  |  |
| 67 | Salicylic acid | -2.20 |  |  | Non-sensitizer |
| 68 | Acetylsalicylic acid | -2.14 | Non-sensitizer |  |  |
| 69 | Methylhydroxybenzoate | -2.04 | Non-sensitizer |  |  |
| 70 | Coumarin | -2.04 |  |  | Non-sensitizer |
| 71 | Methyl 4-hydroxybenzoate | -2.04 |  |  | Non-sensitizer |
| 72 | Ethylene dichloride | -2.00 | Non-sensitizer |  |  |
| 73 | Propylene dichloride | -2.00 | Non-sensitizer |  |  |
| 74 | 2-Heptanone | -2.00 | Sensitizer |  |  |
| 75 | Butyl acrylate | -2.00 |  |  | Sensitizer |
| 76 | 2-Phenylethanol | -1.89 | Sensitizer |  |  |
| 77 | n-Hexanol | -1.89 | Sensitizer |  |  |
| 78 | Hexanoic acid | -1.85 | Non-sensitizer |  |  |
| 79 | Indomethacin | -1.83 | Non-sensitizer |  |  |
| 80 | 3-Methylphenol | -1.81 | Sensitizer |  |  |
| 81 | Trichloromethane | -1.80 | Non-sensitizer |  |  |
| 82 | 2-Methylphenol | -1.80 | Sensitizer |  |  |
| 83 | Diethylether | -1.79 | Non-sensitizer |  |  |
| 84 | Diclofenac | -1.74 | Non-sensitizer |  |  |
| 85 | Hydrocortisone hexanoate | -1.74 | Non-sensitizer |  |  |
| 86 | Heptanoic acid | -1.70 | Non-sensitizer |  |  |
| 87 | Octanoic acid | -1.60 |  |  | Non-sensitizer |
| 88 | 2-Naphthol | -1.55 | Sensitizer |  |  |
| 89 | 1-Naphthol | -1.55 |  |  | Sensitizer |
| 90 | n-Heptanol | -1.50 | Sensitizer |  |  |
| 91 | 4-Ethylphenol | -1.46 | Sensitizer |  |  |
| 92 | 3,4-Xylenol | -1.44 | Sensitizer |  |  |
| 93 | Ibuprofen | -1.44 | Sensitizer |  |  |
| 94 | Chloroxylenol | -1.28 | Sensitizer |  |  |
| 95 | Octanol | -1.28 |  |  | Sensitizer |
| 96 | Thymol | -1.26 | Sensitizer |  |  |
| 97 | 2,4-Dichlorophenol | -1.22 | Sensitizer |  |  |
| 98 | n-Nonanol | -1.22 | Sensitizer |  |  |
| 99 | Hydrocortisone octanoate | -1.21 | Non-sensitizer |  |  |
| 100 | n-Decanol | -1.10 | Sensitizer |  |  |
| 101 | Malathion | -0.69 | Non-sensitizer |  |  |
| 102 | Hydroxyethyl-ethylenediamine |  |  | -3.90 | Sensitizer |
| 103 | Propylene glycol |  |  | -3.79 | Non-sensitizer |
| 104 | N-Methyl-N-nitrosourea |  |  | -3.68 | Sensitizer |
| 105 | N-Ethyl-N-nitrosourea |  |  | -3.61 | Sensitizer |
| 106 | Ethylenediamine |  |  | -3.52 | Sensitizer |
| 107 | Diethylenetriamine |  |  | -3.48 | Sensitizer |
| 108 | 3-Phenylene-diamine |  |  | -3.40 | Sensitizer |
| 109 | Dipropylene triamine |  |  | -3.36 | Sensitizer |
| 110 | 2,5-Diaminotoluene |  |  | -3.31 | Sensitizer |
| 111 | 1,2-Diaminocyclohexane |  |  | -3.29 | Sensitizer |
| 112 | 2,3-Dihydro-1,3-oxazol-2-one |  |  | -3.28 | Sensitizer |
| 113 | 3-Dimethylamino- propylamine |  |  | -3.24 | Sensitizer |
| 114 | 1,4-Dihydroquinone |  |  | -3.23 | Sensitizer |
| 115 | 2-Amino-di- phenylamine |  |  | -3.13 | Sensitizer |
| 116 | Sulfanilamide |  |  | -3.09 | Non-sensitizer |
| 117 | Butanedione |  |  | -3.09 | Sensitizer |
| 118 | 1(2H)-Pyridazinamine |  |  | -3.06 | Non-sensitizer |
| 119 | Glyoxal |  |  | -3.06 | Sensitizer |
| 120 | 3-Aminophenol |  |  | -3.03 | Sensitizer |
| 121 | m-Phenylenebis-(methylamine) |  |  | -3.03 | Sensitizer |
| 122 | Dimethyl carbonate |  |  | -3.02 | Non-sensitizer |
| 123 | 1,3-benzenedimethanamine |  |  | -2.97 | Sensitizer |
| 124 | beta-Propiolactone |  |  | -2.97 | Sensitizer |
| 125 | 5-Amino-2-methylbenzene-sulfonamide |  |  | -2.96 | Non-sensitizer |
| 126 | Sodium ethyl xanthate |  |  | -2.96 | Sensitizer |
| 127 | Dimethyl sulfoxide |  |  | -2.95 | Sensitizer |
| 128 | Methyl methanesulfonate |  |  | -2.94 | Sensitizer |
| 129 | Methyl pyruvate |  |  | -2.89 | Sensitizer |
| 130 | Sulfanilic acid |  |  | -2.88 | Non-sensitizer |
| 131 | 6-Methoxy-4-methyl-8-quinolinamine |  |  | -2.87 | Non-sensitizer |
| 132 | 2,3-Dimethyl-2H-indazol-6-amine |  |  | -2.86 | Non-sensitizer |
| 133 | 5-Amino-O-Cresol |  |  | -2.86 | Sensitizer |
| 134 | 2-Aminophenol |  |  | -2.84 | Sensitizer |
| 135 | 4-Methylamino- phenol sulfate |  |  | -2.82 | Sensitizer |
| 136 | 6-[(2-Methyl-3-pyridinyl)oxy]-3-pyridinamine |  |  | -2.77 | Sensitizer |
| 137 | Glutaraldehyde |  |  | -2.72 | Sensitizer |
| 138 | tert-Butyl 3-Aminobenzoate |  |  | -2.69 | Non-sensitizer |
| 139 | 3-Fluoro-5-(3-pyridinyl)benzen-amine |  |  | -2.64 | Sensitizer |
| 140 | Methyl methacrylate |  |  | -2.63 | Sensitizer |
| 141 | 2,4-Dichloro-pyrimidine |  |  | -2.61 | Sensitizer |
| 142 | 5-Methyl-2,3-hexanedione |  |  | -2.57 | Sensitizer |
| 143 | Anthranilic acid |  |  | -2.53 | Non-sensitizer |
| 144 | (2E)-2-Methyl-2-butenal |  |  | -2.52 | Non-sensitizer |
| 145 | Saccharin |  |  | -2.49 | Non-sensitizer |
| 146 | 5,5-Dimethyl-3-methylenedihydro-2(3H)-furanone |  |  | -2.48 | Sensitizer |
| 147 | Eugenol |  |  | -2.48 | Sensitizer |
| 148 | 2,4-Dinitrobenzene sulfonic acid |  |  | -2.47 | Sensitizer |
| 149 | Isoeugenol |  |  | -2.47 | Sensitizer |
| 150 | 1-Phenyl-1,2-propanedione |  |  | -2.45 | Sensitizer |
| 151 | 2-(3,4-Dimethylphenyl)-5-methyl-2,4-dihydro-3H-pyrazol-3-one |  |  | -2.45 | Sensitizer |
| 152 | 3-Hydroxy-4-methoxybenzaldehyde |  |  | -2.44 | Non-sensitizer |
| 153 | Vanillin |  |  | -2.44 | Non-sensitizer |
| 154 | 2-[1-(4-Bromophenyl)-1-phenylethoxy]-N,N-dimethyl-ethanamine hydrochloride |  |  | -2.44 | Sensitizer |
| 155 | Ethyl vanillin |  |  | -2.43 | Non-sensitizer |
| 156 | Abietic acid |  |  | -2.43 | Sensitizer |
| 157 | Farnesal |  |  | -2.42 | Sensitizer |
| 158 | 4-Hydroxybenzoic acid |  |  | -2.41 | Non-sensitizer |
| 159 | Vinylidene dichloride |  |  | -2.39 | Non-sensitizer |
| 160 | Linalool aldehyde |  |  | -2.39 | Sensitizer |
| 161 | Veratraldehyde |  |  | -2.39 | Sensitizer |
| 162 | Isopropyl myristate |  |  | -2.38 | Sensitizer |
| 163 | Bromobutane |  |  | -2.37 | Non-sensitizer |
| 164 | 3-Hydroxy-2-phenyl-4-quinolinecarboxylic acid |  |  | -2.36 | Non-sensitizer |
| 165 | 3-Bromomethyl-5, 5'-dimethyl- dihydro-2(3H)-furanone |  |  | -2.36 | Sensitizer |
| 166 | 5-Chloro-2-methyl- 4-isothiazolin-3-one |  |  | -2.35 | Sensitizer |
| 167 | Methyl 4-(bromomethyl)benzoate |  |  | -2.32 | Sensitizer |
| 168 | 1-Phenylpropylamine |  |  | -2.31 | Non-sensitizer |
| 169 | Diethylacetaldehyde |  |  | -2.30 | Sensitizer |
| 170 | Vinylpyridine |  |  | -2.30 | Sensitizer |
| 171 | (2-oxo-1-phenylpyrrolidin-3-yl)-Triphenylphosphanium |  |  | -2.29 | Non-sensitizer |
| 172 | 6-Iodo-quinazolin-4-ol |  |  | -2.28 | Non-sensitizer |
| 173 | 2(E)-Hexenal |  |  | -2.28 | Sensitizer |
| 174 | Carvoxime |  |  | -2.28 | Sensitizer |
| 175 | (2E,4E)-2,4-Hexadienal |  |  | -2.27 | Sensitizer |
| 176 | 2-Methoxy-4-methylphenol |  |  | -2.27 | Sensitizer |
| 177 | 2-Acetylcyclohexanone |  |  | -2.26 | Non-sensitizer |
| 178 | 1,2-Benzisothiazolin-3-one |  |  | -2.26 | Sensitizer |
| 179 | Sodium-3,3,5-trimethyl-benzenesulfonate |  |  | -2.24 | Sensitizer |
| 180 | 6-Methoxy-4-methyl-2(1H)-quinolinone |  |  | -2.23 | Non-sensitizer |
| 181 | Methyl salicylate |  |  | -2.23 | Non-sensitizer |
| 182 | 4'-Methoxyacetophenone |  |  | -2.22 | Non-sensitizer |
| 183 | Linoleic acid |  |  | -2.22 | Sensitizer |
| 184 | Linolenic acid |  |  | -2.22 | Sensitizer |
| 185 | Oleic acid |  |  | -2.22 | Sensitizer |
| 186 | 4-Isopropenyl-1-methyl-2-methylene-cyclohexane |  |  | -2.21 | Non-sensitizer |
| 187 | N-(2-Chloro-4-pyrimidinyl)-2,3-dimethyl-2H-indazol-6-amine |  |  | -2.21 | Non-sensitizer |
| 188 | Propylparaben |  |  | -2.20 | Non-sensitizer |
| 189 | 2-(4-tert-Amylcyclohexyl)acetaldehyde |  |  | -2.20 | Sensitizer |
| 190 | 3-Propylidene phthalide |  |  | -2.20 | Sensitizer |
| 191 | Limonene |  |  | -2.20 | Sensitizer |
| 192 | 3-Propoxybenzoic acid |  |  | -2.19 | Non-sensitizer |
| 193 | 2-Methyl-4H-3,1-benzoxazin-4-one |  |  | -2.19 | Sensitizer |
| 194 | Benzoquinone |  |  | -2.18 | Sensitizer |
| 195 | b-Phenylcinnamaldehyde |  |  | -2.17 | Sensitizer |
| 196 | Hydroxycitronellal |  |  | -2.17 | Sensitizer |
| 197 | 4-(Bromomethyl)- benzoic acid ethyl ester |  |  | -2.16 | Sensitizer |
| 198 | 5-(prop-1-en-1-yl)-2-(propan-2-yloxy)phenol |  |  | -2.16 | Sensitizer |
| 199 | 7-Bromotetradecane |  |  | -2.16 | Sensitizer |
| 200 | 6-Diethylaminohexyl bromide hydrobromide |  |  | -2.14 | Sensitizer |
| 201 | Benzyl benzoate |  |  | -2.14 | Sensitizer |
| 202 | b-Phellandrene |  |  | -2.14 | Sensitizer |
| 203 | Palmitoyl chloride |  |  | -2.14 | Sensitizer |
| 204 | 2-Bromo-5-propoxybenzoic acid |  |  | -2.13 | Non-sensitizer |
| 205 | 2-Ethylhexyl acrylate |  |  | -2.12 | Sensitizer |
| 206 | 2-Mercapto-benzothiazole |  |  | -2.12 | Sensitizer |
| 207 | Phenyl benzoate |  |  | -2.12 | Sensitizer |
| 208 | 4-Isopropyl-1-methylene-cyclohexane |  |  | -2.11 | Non-sensitizer |
| 209 | 4-Chloro-6- iodoquinazoline |  |  | -2.11 | Sensitizer |
| 210 | Bromoheptadecane |  |  | -2.11 | Sensitizer |
| 211 | Bromohexadecane |  |  | -2.11 | Sensitizer |
| 212 | Bromopentadecane |  |  | -2.11 | Sensitizer |
| 213 | Clorohexadecane |  |  | -2.11 | Sensitizer |
| 214 | Norbornene fluoroalcohol |  |  | -2.11 | Sensitizer |
| 215 | Bromooctadecane |  |  | -2.10 | Sensitizer |
| 216 | 1-Chlorooctadecane |  |  | -2.09 | Sensitizer |
| 217 | 2,4-Heptadienal |  |  | -2.09 | Sensitizer |
| 218 | Bromoeicosane |  |  | -2.09 | Sensitizer |
| 219 | Benzaldehyde |  |  | -2.06 | Non-sensitizer |
| 220 | 12-Bromododecanoic acid |  |  | -2.06 | Sensitizer |
| 221 | Bromodocosane |  |  | -2.06 | Sensitizer |
| 222 | Bromotetradecane |  |  | -2.05 | Sensitizer |
| 223 | Chlorotetradecane |  |  | -2.05 | Sensitizer |
| 224 | 6-Methylcoumarin |  |  | -2.04 | Non-sensitizer |
| 225 | 2-Bromo-5-hydroxy-benzaldehyde |  |  | -2.04 | Sensitizer |
| 226 | Ethyl 4-iodobenzoate |  |  | -2.04 | Sensitizer |
| 227 | 1-Chloromethyl- pyrene |  |  | -2.02 | Sensitizer |
| 228 | Perillaldehyde |  |  | -2.02 | Sensitizer |
| 229 | 2,6,6-Trimethyl-1,3-cyclohexadiene-1-carbaldehyde |  |  | -1.99 | Sensitizer |
| 230 | alpha-Hexyl cinnamic aldehyde |  |  | -1.99 | Sensitizer |
| 231 | R-Carvone |  |  | -1.99 | Sensitizer |
| 232 | Quinoxyfen |  |  | -1.98 | Non-sensitizer |
| 233 | Bromohexane |  |  | -1.98 | Sensitizer |
| 234 | 3,4-Dihydrocoumarin |  |  | -1.96 | Sensitizer |
| 235 | 4-Nitrobenzyl bromide |  |  | -1.94 | Sensitizer |
| 236 | 2-Chloro-6-methoxy-4-methylquinoline |  |  | -1.93 | Non-sensitizer |
| 237 | Phenylacetaldehyde |  |  | -1.93 | Sensitizer |
| 238 | Undecylenic acid |  |  | -1.92 | Sensitizer |
| 239 | 1-(p-methoxyphenyl)-1- penten-3-one |  |  | -1.91 | Sensitizer |
| 240 | Citral |  |  | -1.90 | Sensitizer |
| 241 | 2-methoxy-5-propylphenol |  |  | -1.89 | Sensitizer |
| 242 | 4-Allylanisole |  |  | -1.89 | Sensitizer |
| 243 | 5-Methyl-2-phenyl-2-hexenal |  |  | -1.89 | Sensitizer |
| 244 | alpha-Amyl cinnamic aldehyde |  |  | -1.89 | Sensitizer |
| 245 | Dihydroeugenol |  |  | -1.89 | Sensitizer |
| 246 | Undec-10-enal |  |  | -1.87 | Sensitizer |
| 247 | 3-Chloro-4-fluorobenzoylchloride |  |  | -1.86 | Sensitizer |
| 248 | 4-isopropenyl-1-methyl-6-methylene-cyclohexene |  |  | -1.84 | Sensitizer |
| 249 | Cinnamic aldehyde |  |  | -1.84 | Sensitizer |
| 250 | Iodooctadecane |  |  | -1.83 | Non-sensitizer |
| 251 | alpha-Butyl cinnamic aldehyde |  |  | -1.83 | Sensitizer |
| 252 | (E)-Anethol |  |  | -1.82 | Sensitizer |
| 253 | Cinnamic alcohol |  |  | -1.82 | Sensitizer |
| 254 | beta-Terpinene |  |  | -1.81 | Non-sensitizer |
| 255 | 6(E)-Nonenal |  |  | -1.81 | Sensitizer |
| 256 | 6-Chloro-1-hexanol |  |  | -1.80 | Non-sensitizer |
| 257 | 3,5,5-Trimethyl- hexanoyl chloride |  |  | -1.80 | Sensitizer |
| 258 | (4Z)-2-Methyl-6-methylene-4-octene |  |  | -1.79 | Non-sensitizer |
| 259 | p-tert-Butyl-alpha-ethyl-hydrocinnamal |  |  | -1.78 | Sensitizer |
| 260 | Isononanoyl chloride |  |  | -1.77 | Sensitizer |
| 261 | p-Methylhydrocinnamic aldehyde |  |  | -1.77 | Sensitizer |
| 262 | m-Chloropropiophenone |  |  | -1.76 | Non-sensitizer |
| 263 | (2E)-2-Decenal |  |  | -1.76 | Sensitizer |
| 264 | alpha-Methyl cinnamic aldehyde |  |  | -1.76 | Sensitizer |
| 265 | alpha-Methyl phenylacetaldehyde |  |  | -1.76 | Sensitizer |
| 266 | Hexane |  |  | -1.75 | Non-sensitizer |
| 267 | Cyclamen aldehyde |  |  | -1.74 | Sensitizer |
| 268 | Bromotridecane |  |  | -1.73 | Sensitizer |
| 269 | Geraniol |  |  | -1.73 | Sensitizer |
| 270 | Chlorobenzene |  |  | -1.72 | Non-sensitizer |
| 271 | Benzylidene acetone |  |  | -1.72 | Sensitizer |
| 272 | Nonanoyl chloride |  |  | -1.72 | Sensitizer |
| 273 | Bromododecane |  |  | -1.71 | Sensitizer |
| 274 | 2-Methylundecanal |  |  | -1.70 | Sensitizer |
| 275 | Iodohexane |  |  | -1.69 | Non-sensitizer |
| 276 | Bromoundecane |  |  | -1.67 | Sensitizer |
| 277 | Benzyl Bromide |  |  | -1.64 | Sensitizer |
| 278 | alpha-Terpinene |  |  | -1.62 | Sensitizer |
| 279 | Bromononane |  |  | -1.60 | Non-sensitizer |
| 280 | Chlorononane |  |  | -1.58 | Non-sensitizer |
| 281 | 1-Iododecane |  |  | -1.47 | Sensitizer |
| 282 | Iodononane |  |  | -1.45 | Sensitizer |
| 283 | Iododecane |  |  | -1.41 | Sensitizer |

Notes: Kp: permeability coefficient.

# Table S6. Compounds within Cluster 1 of united skin sensitization and permeability dataset.

| **Compound name** | **logKp** | **Predicted**  **logKp** | **Sensitization**  **Activity** | **Predicted**  **Sensitization** |
| --- | --- | --- | --- | --- |
| Atropine | -5.07 |  |  | Non-sensitizer |
| Scopolamine | -4.30 |  |  | Non-sensitizer |
| Barbital | -3.95 |  |  | Non-sensitizer |
| Butobarbital | -3.71 |  |  | Non-sensitizer |
| Phenobarbital | -3.34 |  |  | Non-sensitizer |
| Griseofulvin | -2.89 |  |  | Non-sensitizer |
| Amylobarbital | -2.64 |  |  | Non-sensitizer |
| Ibuprofen | -1.44 |  |  | Sensitizer |
| Sodium-3,3,5-trimethyl-benzenesulfonate |  | -2.24 | Sensitizer |  |
| p-tert-Butyl-a-ethyl-hydrocinnamal |  | -1.78 | Sensitizer |  |
| Cyclamen aldehyde |  | -1.74 | Sensitizer |  |

Notes: 1: sensitizer; 0: non-sensitizer; Kp: permeability coefficient.

# Figure S1. Skin permeability *vs.* sensitization imputed data for united dataset, containing 174 sensitizers and 109 non-sensitizers.

Notes: Kp: permeability coefficient.

# Figure S2. Skin permeability *vs.* sensitization imputed data for the Cluster 1 of united dataset, containing 4 sensitizers and 7 non-sensitizers.

Notes: Kp: permeability coefficient.

# REFERENCES

(1) Consonni, V. et al. *J. Chem. Inf. Model.* **2009**, *49*, 1669–1678.

(2) Consonni, V. et al. *J. Chemom.* **2010**, *24*, 194–201.

(3) Roy, K. et al. *Comb. Chem. High Throughput Screen.* **2011**, *14*, 450–474.

(4) Ojha, P. K. et al. *Chemom. Intell. Lab. Syst.* **2011**, *107*, 194–205.

(5) Roy, K. et al. *J. Chem. Inf. Model.* **2012**, *52*, 396–408.

(6) Roy, K. et al. *Mini Rev. Med. Chem.* **2012**, *12*, 491–504.

(7) Flynn, G. L. In *Principles of Route-to-Route Extrapolation for Risk assessment*; Gerrity, T. R., Ed.; Elsevier: New York, NY, 1990; pp. 93–117.

(8) Potts, R. O. et al. *Pharm. Res.* **1992**, *9*, 663–669.

(9) Scheuplein, R. J. et al. *J. Invest. Dermatol.* **1973**, *60*, 286–296.

(10) Scheuplein, R. J. et al. In *Biochemistry and physiology of the skin*; Goldsmith, L. A., Ed.; Oxford University Press: New York, NY, 1983; pp. 1255–1295.

(11) Scheuplein, R. J. et al. *J. Invest. Dermatol.* **1969**, *52*, 63–70.

(12) Lien, E. J. et al. *Pharm. Res.* **1995**, *12*, 583–587.

(13) Barratt, M. D. *Toxicol. In Vitro* **1995**, *9*, 27–37.

(14) Potts, R. O. et al. *Pharm. Res.* **1995**, *12*, 1628–1633.

(15) Hostýnek, J. J. et al. *Quant. Struct. relationships* **1997**, *16*, 473–479.

(16) Johnson, M. E. et al. *J. Pharm. Sci.* **1995**, *84*, 1144–1146.

(17) Abraham, M. H. et al. *Pestic. Sci.* **1999**, *55*, 78–88.

(18) Kirchner, L. A. et al. *ATLA* **1997**, *25*, 359–370.

(19) Cronin, M. T. et al. *Eur. J. Pharm. Sci.* **1999**, *7*, 325–330.

(20) Degim, I. T. et al. *Int. J. Pharm.* **1998**, *170*, 129–133.

(21) Moss, G. P. et al. *Int. J. Pharm. Pharm.* *238*, 105–109.

(22) Wilschut, A. et al. *Chemosphere* **1995**, *30*, 1275–1296.

(23) Patel, H. et al. *Chemosphere* **2002**, *48*, 603–613.

(24) Magnusson, B. M. et al. *J. Invest. Dermatol.* **2004**, *122*, 993–999.

(25) Vecchia, B. E. et al. In *Transdermal Drug Delivery*; Guy, R. H., Ed.; Marcel Dekker, 2002; pp. 57–141.

(26) Magnusson, B. M. et al. *J. Invest. Dermatol.* **2004**, *122*, 993–999.

(27) Berge, W. *Chemosphere* **2009**, *75*, 1440–1445.

(28) Johnson, M. E. et al. *J. Pharm. Sci.* **1997**, *86*, 1162–1172.

(29) Chen, L. L. et al. *Acta Pharmacol. Sin.* **2007**, *28*, 591–600.

(30) Chen, L. et al. *AIChE J.* **2010**, *56*, 1136–1146.

(31) Kielhorn, J. et al. *The Dermal Absorption*; Fraunhofer Institute Toxicology and Experimental Medicine: Hanover, 2005; pp. 90–91.

(32) Chauhan, P. et al. *SAR QSAR Environ. Res.* **2010**, *21*, 481–494.

(33) Moss, G. P. et al. *J. Pharm. Pharmacol.* **2011**, *63*, 1411–1427.

(34) Riviere, J. E. et al. *Toxicol. Sci.* **2011**, *119*, 224–232.

(35) Khajeh, A. et al. *SAR QSAR Environ. Res.* **2014**, *25*, 35–50.